

# Cascading compliance requirements for thesis-related data and software

A survey of institutional policy at American (AAU) research  
universities

**Gail Peretsman-Clement**

19th International Symposium on Electronic Theses and Dissertations  
Lille, France

July 11, 2016

**Caltech** Library

- A. Introduction: Why Worry About Compliance Requirements for Thesis Data?

- A. Introduction: Why Worry About Compliance Requirements for Thesis Data?
- B. Compliance Requirements Impact Legal Interoperability of Thesis Data

- A. Introduction: Why Worry About Compliance Requirements for Thesis Data?
- B. Compliance Requirements Impact Legal Interoperability of Thesis Data
- C. Research Question: Do University IP Policies Adhere to Legal Interoperability Principles for Thesis Data??

- A. Introduction: Why Worry About Compliance Requirements for Thesis Data?
- B. Compliance Requirements Impact Legal Interoperability of Thesis Data
- C. Research Question: Do University IP Policies Adhere to Legal Interoperability Principles for Thesis Data??
- D. Recommendations and Next Steps

# Why Worry About Compliance for Thesis Data?

[Home](#) > [Research](#)

## Caltech

[About Caltech](#)

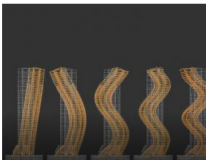
[News & Events](#)

[Research & Education](#)

[Join Us](#)

### RESEARCH

Top-caliber scientists and engineers lead breakthrough research that shifts paradigms, launches new fields, creates technologies, and benefits society.



[Research News](#) >



[Faculty](#) >



[Technology Transfer and Corporate Partnerships](#)



[Student Research Programs](#) >



[Research Centers](#) >

**NEWS** • [Community Seismic Network Detected Air Pulse From Refinery Explosion](#)

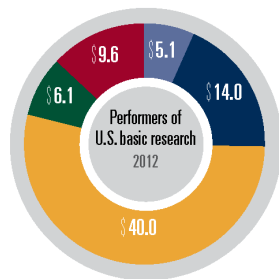
CALIFORNIA INSTITUTE OF TECHNOLOGY  
Site content Copyright © 2016 California Institute of Technology

1200 EAST CALIFORNIA BOULEVARD, PASADENA, CALIFORNIA 91125  
[Contact Us](#) [Report a Copyright Infringement](#) [Privacy Statement](#)



# Graduate student researchers are the **workhorses** of the American research enterprise

FIGURE 3 **PERFORMERS OF U.S. BASIC RESEARCH (2012)**



SOURCE: National Science Foundation, National Center for Science and Engineering Statistics. 2013. National Patterns of R&D Resources: 2011–12 Data Update. Dollars are in billions. Total basic research funding for 2012 was \$74,849 in current dollars.

Control of institution and primary function/occupational activity	All staff	
	Total	Full-time
<b>Total staff</b>	<b>3,920,836</b>	<b>2,489,533</b>
Staff whose primary responsibility is instruction, research, and/or public service	1,565,504	784,639
Primarily instruction	1,195,544	499,080
Instruction combined with research and/or public service	273,222	208,473
Primarily research	71,357	59,953
Primarily public service <sup>2</sup>	25,381	17,133
Executive/administrative/managerial	248,982	241,136
Other professional (support/service)	816,166	708,924
Graduate assistants <sup>3</sup>	355,916	↑
Technical and paraprofessional	199,538	158,473
Clerical and secretarial	436,018	340,600
Skilled crafts	61,130	58,448
Service/maintenance	237,582	197,313

Figure: NCES 2012, Table 1, p. 9

Figure: AAU 2015, Fig. 3

# IP around Graduate Student research tends to be fuzzy because of multiple interests and stakeholders

- 1 Graduate research is commonly **supported with sponsor funds** and thus subject to **compliance** requirements around sharing of research results
- 2 Graduate students are 'sub-faculty' : **employees with less agency** of their work than faculty or staff scientists
- 3 Graduate students in research-intensive disciplines publish their findings in **multiple disparate channels** that need to be **linked together**



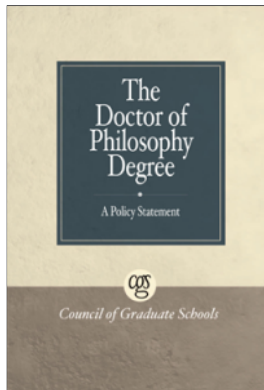
# To Achieve Greatest Possible Data Sharing, We Need to Carefully Curate Graduate Student Data, \*particularly\*



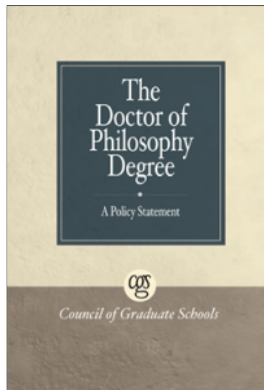
Figure: Sarah Pease carefully shucks oysters to avoid damaging the animal inside, Virginia Sea Grant, 2008, <https://www.flickr.com/photos/virginiaseagrant/21073901051>, (CC BY-ND 2.0)

# Legal Status of the Thesis vs. Thesis Data

- The doctoral candidate retains copyright over the dissertation.

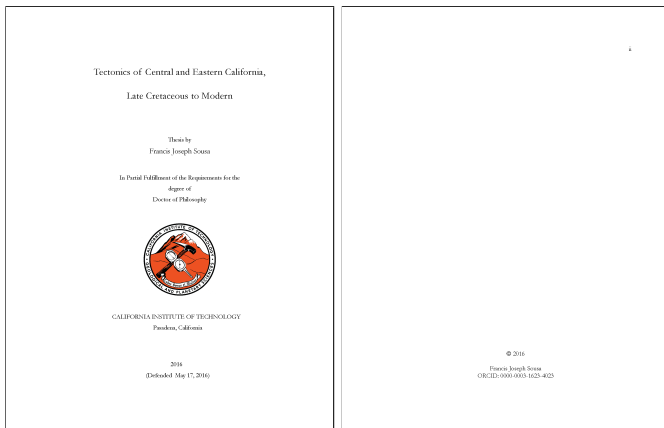


# Legal Status of the Thesis vs. Thesis Data




- The doctoral candidate retains copyright over the dissertation.
- The university typically retains ownership of any research products described within the dissertation, or otherwise produced during candidacy that result from the student's use of university resources (including financial support), facilities, and faculty supervision. In the period during which advisers and advisees attempt to find a productive match, doctoral candidates may be asked to sign institutional agreements regarding how authorship responsibilities for research results will be allocated...

# Thesis Sharing *not* = Data Sharing



# Thesis Data = Complex Set of Objects linked across Disparate Platforms

<b>Funders:</b>	<table><tr><th>Funding Agency</th><th>Grant Number</th></tr><tr><td>Gordon and Betty Moore Foundation</td><td>GBMF #423.01</td></tr><tr><td>National Science Foundation</td><td>EAR-0408526</td></tr><tr><td>Southern California Earthquake Center</td><td>12188, 13151, 14160</td></tr><tr><td>NSF Cooperative Agreement</td><td>EAR-1033462</td></tr><tr><td>USGS Cooperative Agreement</td><td>G12AC20038</td></tr></table>	Funding Agency	Grant Number	Gordon and Betty Moore Foundation	GBMF #423.01	National Science Foundation	EAR-0408526	Southern California Earthquake Center	12188, 13151, 14160	NSF Cooperative Agreement	EAR-1033462	USGS Cooperative Agreement	G12AC20038
Funding Agency	Grant Number												
Gordon and Betty Moore Foundation	GBMF #423.01												
National Science Foundation	EAR-0408526												
Southern California Earthquake Center	12188, 13151, 14160												
NSF Cooperative Agreement	EAR-1033462												
USGS Cooperative Agreement	G12AC20038												
<b>Record Number:</b>	CaltechTHESIS:05252016-090108907												
<b>Persistent URL:</b>	<a href="http://resolver.caltech.edu/CaltechTHESIS:05252016-090108907">http://resolver.caltech.edu/CaltechTHESIS:05252016-090108907</a>												
<b>DOI:</b>	10.7907/Z9H41PD8												
<b>Related URLs:</b>	<table><tr><th>URL</th><th>URL Type</th><th>Description</th></tr><tr><td><a href="http://dx.doi.org/10.1016/j.epsl.2016.01.020">http://dx.doi.org/10.1016/j.epsl.2016.01.020</a></td><td>DOI</td><td>Article adapted for ch. 1</td></tr><tr><td><a href="http://dx.doi.org/10.7907/Z9CJ8BF4">http://dx.doi.org/10.7907/Z9CJ8BF4</a></td><td>DOI</td><td>Dataset for 2014 Hector Mine field research</td></tr></table>	URL	URL Type	Description	<a href="http://dx.doi.org/10.1016/j.epsl.2016.01.020">http://dx.doi.org/10.1016/j.epsl.2016.01.020</a>	DOI	Article adapted for ch. 1	<a href="http://dx.doi.org/10.7907/Z9CJ8BF4">http://dx.doi.org/10.7907/Z9CJ8BF4</a>	DOI	Dataset for 2014 Hector Mine field research			
URL	URL Type	Description											
<a href="http://dx.doi.org/10.1016/j.epsl.2016.01.020">http://dx.doi.org/10.1016/j.epsl.2016.01.020</a>	DOI	Article adapted for ch. 1											
<a href="http://dx.doi.org/10.7907/Z9CJ8BF4">http://dx.doi.org/10.7907/Z9CJ8BF4</a>	DOI	Dataset for 2014 Hector Mine field research											
<b>ORCID:</b>	<table><tr><th>Author</th><th>ORCID</th></tr><tr><td>Sousa, Francis Joseph</td><td><a href="https://orcid.org/0000-0003-1623-4023">0000-0003-1623-4023</a></td></tr></table>	Author	ORCID	Sousa, Francis Joseph	<a href="https://orcid.org/0000-0003-1623-4023">0000-0003-1623-4023</a>								
Author	ORCID												
Sousa, Francis Joseph	<a href="https://orcid.org/0000-0003-1623-4023">0000-0003-1623-4023</a>												
<b>Default Usage Policy:</b>	No commercial reproduction, distribution, display or performance rights in this work are provided.												
<b>ID Code:</b>	9748												
<b>Collection:</b>	CaltechTHESIS												
<b>Deposited By:</b>	Francis Sousa												
<b>Deposited On:</b>	25 May 2016 23:34												
<b>Last Modified:</b>	25 May 2016 23:55												
<b>Thesis Files</b>													
 <a href="#">PDF (PhD Dissertation) - Final Version</a> Restricted to Caltech community only until 31 December 2016													

# Thesis Data Sharing:

## Multiple ownership interests and rights assertions

### Other Contributors:

Contribution	Other Contributors Name	Identifier	URI (may be blank)
Thesis advisor	Stock, Joann M.	Stock-J-M	UNSPECIFIED
Research team member	Scharer, Katherine	Scharer-K	UNSPECIFIED
Research team member	Hudnut, Kenneth	Hudnut-K-W	UNSPECIFIED
Research team member	Akciz, S. O.	Akciz-S-O	UNSPECIFIED
Research team member	Witkosky, Ryan D.	Witkosky-R-D	UNSPECIFIED
Research team member	Harvey, Janet C.	Harvey-J-C	UNSPECIFIED

**Additional Information:** This research was supported by the Southern California Earthquake Center (Contribution No. 6121). SCEC is funded by NSF Cooperative Agreement EAR-1033462 & USGS Cooperative Agreement G12AC20038.

**Group:** Seismological Laboratory

### Funders:

Funding Agency	Grant Number
Southern California Earthquake Center (SCEC)	
NSF	EAR-1033462
USGS	G12AC20038

**Subject Keywords:** Earthquake geology; Faults; Strike-slip

### Other Numbering System:

Other Numbering System Name	Other Numbering System ID
Southern California Earthquake Center	6121

**DOI:** 10.7907/Z9CJ8BF4

**Record Number:** CaltechAUTHORS:20160509-112713434

**Persistent URL:** <http://resolver.caltech.edu/CaltechAUTHORS:20160509-112713434>

### Usage Policy:

Copyright (c) 2016, California Institute of Technology. All rights not granted herein are expressly reserved by Caltech. Redistribution and use, with or without modification, are permitted provided that the following conditions are met: 1. Redistributions must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. 2. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this dataset without specific prior written permission. THIS DATASET IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR

# Thesis Data Sharing:

## Multiple ownership interests and rights assertions



### Copyright and Software Policy

Copyrights to and royalties from textbooks, reference works, submissions to scientific journals, and other copyrightable materials (except for computer software, which is treated below) produced by Faculty members as a part of their normal teaching and scholarly activities at the Institute that do not result from projects specifically funded in whole or in part by the Institute or by a sponsor of the Institute, shall belong to the author or authors and may be retained by them. If, on the other hand, the Institute provides funds, or a sponsor's funds, to finance (in whole or in part) a specific research or educational project and copyrightable materials are produced by employees as a result of the project, the ownership of copyrights and royalty rights therein shall be held by the Institute.]

All rights to computer software, including computer programs, computer databases, and associated documentation whether copyrightable or patentable, produced by employees or students in the line of Institute duty or with any use of Institute facilities, shall be owned by the Institute, regardless of the source of funds used to produce the computer software.

It is the policy of the Institute that such software is used for the public benefit, and the Institute intends to serve the public interest by prudent and appropriate efforts to transfer the technology to those who will facilitate public use. When this result is achieved by the licensing of computer

# Thesis Data Sharing Governed by Cascading Compliance Requirements

## 1 Copyright Statutes



# Thesis Data Sharing Governed by Cascading Compliance Requirements

- 1 Copyright Statutes
- 2 Other governing legal doctrines

# Thesis Data Sharing Governed by Cascading Compliance Requirements

- 1 Copyright Statutes
- 2 Other governing legal doctrines
- 3 Sponsor policy

# Thesis Data Sharing Governed by Cascading Compliance Requirements

- 1 Copyright Statutes
- 2 Other governing legal doctrines
- 3 Sponsor policy
- 4 Other contractual terms and conditions

# Thesis Data Sharing Governed by Cascading Compliance Requirements

- 1 Copyright Statutes
- 2 Other governing legal doctrines
- 3 Sponsor policy
- 4 Other contractual terms and conditions
- 5 Institute Intellectual Property Policy

# Thesis Data Sharing Governed by Cascading Compliance Requirements

- 1 Copyright Statutes
- 2 Other governing legal doctrines
- 3 Sponsor policy
- 4 Other contractual terms and conditions
- 5 Institute Intellectual Property Policy
- 6 Normative practices (Disciplinary or Laboratory)

## B. IP Policy (Compliance Requirements) Constrain Legal Interoperability of Thesis Data

### What is Legal Interoperability?



### Legal Interoperability occurs among datasets when:

- Use conditions are clearly and readily determinable for each dataset
- Legal use conditions imposed on each dataset allow creation and use of combined / derivative products
- Users may legally access and use each dataset without need for case-by-case authorization

# CODATA/RDA Principles for Legal Interoperability of Research Data



The simplest cases of legal interoperability occur when datasets are affirmatively identified as having no legal restrictions



# How to Achieve Legal Interoperability of Research Data?



Adhere to Six Basic Principles:

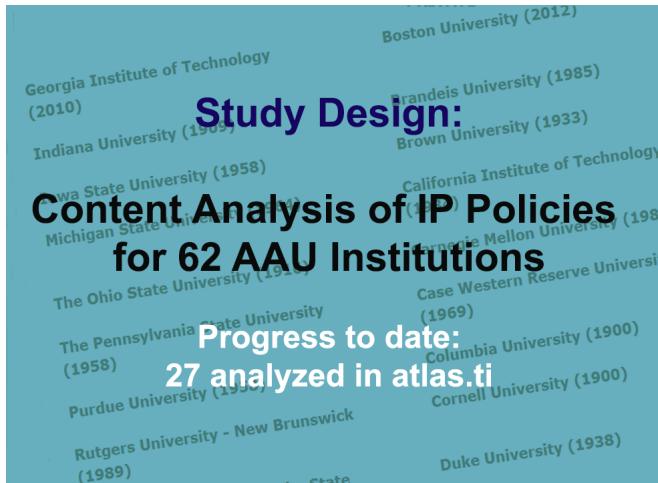
- 1** Facilitate the lawful access to and reuse of research data
- 2** Determine the rights to and responsibilities for the data
- 3** Balance the legal interests
- 4** State the rights transparently and clearly
- 5** Promote the harmonization of rights in research data
- 6** Provide proper attribution and credit for data contributions



# Research Question: Do University IP Policies Adhere to Legal Interoperability Principles for Thesis Data?

**Study Design:**  
**Content Analysis of IP Policies  
for 62 AAU Institutions**

Progress to date:  
27 analyzed in atlas.ti



The background image is a collage of university names and years, representing the 62 AAU institutions studied. The text is tilted and overlapping, creating a sense of a large dataset. Visible institutions include Georgia Institute of Technology (2010), Indiana University (1909), Iowa State University (1958), Michigan State University (1958), The Ohio State University (1958), The Pennsylvania State University (1958), Purdue University (1958), Rutgers University - New Brunswick (1989), Boston University (2012), Brandeis University (1985), Brown University (1933), California Institute of Technology (1933), Carnegie Mellon University (1985), Case Western Reserve University (1969), Columbia University (1900), Cornell University (1900), and Duke University (1938).

# Preliminary Results



Figure: No Negative Cards, geralt,  
<https://pixabay.com/en/no-negative-cards-spiral-68481/> (CC0 Public Domain)

# Conditions based on Student Employment Status

## Student Writings (University of Washington)

"Students employed by the University in any capacity are covered by the terms of this policy. In addition, where a student receives financial aid or remuneration under a sponsored research, training, or fellowship program, his or her rights in copyrightable materials are limited by the terms of the University agreement with the sponsoring agency. The University has no ownership rights in copyrightable materials developed by students who are not employees of the University or in materials unrelated to their employment."

*–Stanford University Research Policy Handbook, 9.2 Copyright Policy*

# Conditions based on Sponsorship of Student Research Project

## Student Works (University of Arizona)

Note that many students are employed or paid by the University in some capacity, often as graduate students under a sponsored research project or department, or faculty-controlled funding. Any such students will be treated the same way as faculty and other University employees for the purposes of this Policy and the ABOR IP Policy with respect to the paid work they perform.

# Conditions based on Use of University Resources

## Student Works-Dissertations (Stanford)

### A. Books, Articles and Similar Works, Including Unpatentable Software

Stanford does not claim ownership to pedagogical, scholarly, or artistic works, regardless of their form of expression. Such works include those of students created in the course of their education, such as dissertations, papers and articles. The University claims no ownership of ...unpatentable software, or other works ...which are not institutional works and did not make significant use of University resources or the services of University non-faculty employees working within the scope of their employment.

## D.Recommendations and Next Steps

- 1 Complete the study to compile thorough dataset
- 2 Share data and findings broadly
  - Publish in OA refereed publication with data publishing policy!
  - Share with University Administrators (who might mobilize through professional associations e.g., NACUA, AAU)
- 3 Use data to promote importance of legal interoperability for thesis data sharing to Graduate School, Faculty Advisors, Graduate Students
- 4 Devise model IP Policy language through professional associations e.g., NDLTD, USETDA, CODATA/RDA

# References

International Council for Science, Committee on Data for Science and Technology and Research Data Alliance (2016). *Legal Interoperability Principles and Guidelines*, <https://rd-alliance.org/groups/rdacodata-legal-interoperability-ig.html>

Association of American Universities (2015). *AAU Policy Brief*, February 2015, No.1, <https://www.aau.edu/WorkArea/DownloadAsset.aspx?id=15974>

National Center for Education Statistics (2012). *Employees in Postsecondary Institutions, Fall 2011* NCES 2012-156rev, Washington DC: US Dept. Education, 6 pp. + Appendix, <http://nces.ed.gov/pubs2012/2012156rev.pdf>

Council of Graduate Schools (2005). *Doctor of Philosophy Degree: A Policy Statement*, Washington D.C., Council of Graduate schools, 36 pp.